

Corruption of parameter behavior and regionalization by model and forcing data errors: A Bayesian example using the SNOW17 model

Minxue He,^{1,2,3} Terri S. Hogue,¹ Kristie J. Franz,⁴ Steven A. Margulis,¹ and Jasper A. Vrugt^{5,6,7}

Received 12 July 2010; revised 20 May 2011; accepted 16 June 2011; published 26 July 2011.

[1] The current study evaluates the impacts of various sources of uncertainty involved in hydrologic modeling on parameter behavior and regionalization utilizing different Bayesian likelihood functions and the Differential Evolution Adaptive Metropolis (DREAM) algorithm. The developed likelihood functions differ in their underlying assumptions and treatment of error sources. We apply the developed method to a snow accumulation and ablation model (National Weather Service SNOW17) and generate parameter ensembles to predict snow water equivalent (SWE). Observational data include precipitation and air temperature forcing along with SWE measurements from 24 sites with diverse hydroclimatic characteristics. A multiple linear regression model is used to construct regionalization relationships between model parameters and site characteristics. Results indicate that model structural uncertainty has the largest influence on SNOW17 parameter behavior. Precipitation uncertainty is the second largest source of uncertainty, showing greater impact at wetter sites. Measurement uncertainty in SWE tends to have little impact on the final model parameters and resulting SWE predictions. Considering all sources of uncertainty, parameters related to air temperature and snowfall fraction exhibit the strongest correlations to site characteristics. Parameters related to the length of the melting period also show high correlation to site characteristics. Finally, model structural uncertainty and precipitation uncertainty dramatically alter parameter regionalization relationships in comparison to cases where only uncertainty in model parameters or output measurements is considered. Our results demonstrate that accurate treatment of forcing, parameter, model structural, and calibration data errors is critical for deriving robust regionalization relationships.

Citation: He, M., T. S. Hogue, K. J. Franz, S. A. Margulis, and J. A. Vrugt (2011), Corruption of parameter behavior and regionalization by model and forcing data errors: A Bayesian example using the SNOW17 model, *Water Resour. Res.*, 47, W07546, doi:10.1029/2010WR009753.

1. Introduction

[2] Snow plays a critical role in the hydrologic cycle in terms of both water supply and regional and global surface energy budgets [e.g., *Cayan*, 1996; *Andreadis and Lettenmaier*, 2006; *Durand and Margulis*, 2006]. There is strong evidence of snowpack decline and altered melt timing in response to atmospheric warming in the western United States [*Cayan et al.*, 2001; *Mote*, 2003; *Service*, 2004; *Stewart et al.*, 2004; *Mote et al.*, 2005; *Kapnick and Hall*,

2010]. Changes in snowpack depth and melt patterns make accurate prediction of snowmelt increasingly critical, particularly in light of a changing climate [*Franz et al.*, 2008]. Snow models are generally developed and applied to simulate snow accumulation and ablation processes and subsequently predict snowmelt. There have been numerous snow modeling studies simulating the mass and energy evolution and thus the snow water equivalent (SWE) of the snowpack via different models (e.g., SNOW17 of *Anderson* [1973], UEB of *Tarboton and Luce* [1996], SAST of *Jin et al.* [1999], ESCIMO of *Strasser et al.* [2002], and SNOWCAN of *Tribbeck et al.* [2004]).

[3] There are a range of snow models with varying complexities. However, successful application relies heavily on the accuracy of associated model parameters [e.g., *Franz et al.*, 2008]. Many parameters cannot be determined directly from the field and can only be derived from calibration against in situ or remotely sensed snow measurements. In the past several decades, a variety of parameter estimation methods have been developed and reported in the hydrologic literature. These methods include the maximum likelihood algorithm [*Restrepo and Bras*, 1985], the generalized likelihood uncertainty estimation (GLUE) [*Beven and Binley*, 1992], the shuffled complex evolution (SCE-UA)

¹Department of Civil and Environmental Engineering, University of California, Los Angeles, California, USA.

²Now at Office of Hydrologic Development, National Weather Service, NOAA, Silver Spring, Maryland, USA.

³Also at Riverside Technology, Inc., Fort Collins, Colorado, USA.

⁴Department of Geological and Atmospheric Sciences, Iowa State University, Ames, Iowa, USA.

⁵Department of Civil and Environmental Engineering, University of California, Irvine, California, USA.

⁶Earth and Environmental Sciences Division, Los Alamos National Laboratory, Los Alamos, New Mexico, USA.

⁷Institute for Biodiversity and Ecosystem Dynamics, University of Amsterdam, Amsterdam, Netherlands.

algorithm [Duan *et al.*, 1992], the multistep automatic calibration scheme [Hogue *et al.*, 2000, 2006b], the Bayesian recursive estimation (BARE) [Thiemann *et al.*, 2001], the Parameter Identification Methods Based on the Localization of Information (PIMLI) [Vrugt *et al.*, 2002], the Dynamic Identifiability Analysis (DYNIA) [Wagener *et al.*, 2003], and Markov chain Monte Carlo (MCMC) methods, including the random walk Metropolis algorithm [Kuczera and Parent, 1998], the shuffled complex evolution Metropolis (SCEM-UA) algorithm [Vrugt *et al.*, 2003], and the Differential Evolution Adaptive Metropolis (DREAM) method [Vrugt *et al.*, 2008a, 2008b]. Most of these methods attribute the mismatch between model predictions and calibration data to parameter uncertainty, without recourse for considering other sources of error. These error sources, including uncertainty associated with forcing (input) data, calibration observations, and model structure (which includes the numerical time-stepping schemes applied to solve governing model equations) [Kitanidis and Bras, 1980; Beck, 1987; Muleta and Nicklow, 2005; Beven, 2006; Gupta *et al.*, 2006; Clark and Kavetski, 2010; Kavetski and Clark, 2010; Schoups *et al.*, 2010], are difficult to determine explicitly in practice, particularly because there is no widely agreed upon method to characterize them [Liu and Gupta, 2007; Montanari, 2007; Gotzinger and Bardossy, 2008; Vrugt *et al.*, 2008a]. A few methods have recently emerged to investigate model forcing and/or structural uncertainty. These methods include the Bayesian model averaging (BMA) [Duan *et al.*, 2007; Vrugt and Robinson, 2007; Franz *et al.*, 2010], the integrated Bayesian uncertainty estimator (IBUNE) [Ajami *et al.*, 2007], the Bayesian total error analysis (BATEA) and closely related developments [Kavetski *et al.*, 2002, 2006; Kuczera *et al.*, 2006; Thyer *et al.*, 2009; Renard *et al.*, 2010], the Simultaneous Optimization and Data Assimilation (SODA) [Vrugt *et al.*, 2005], and the Framework for Understanding Structural Errors (FUSE) [Clark *et al.*, 2008; Clark and Kavetski, 2010]. However, these algorithms have not been commonly applied in practice, partly because of their computational requirements.

[4] Regionalization provides an alternative approach to parameter estimation in which neighboring catchments that exhibit similar geology, soil, vegetation, and climate are used to transfer parameters [e.g., Merz and Bloschl, 2004; Wagener and Wheeler, 2006; Oudin *et al.*, 2008; Buytaert and Beven, 2009]. Regression relationships between catchment characteristics and snow model parameters can also be developed. Such relationships can be built for gauged catchments and used to derive the parameters for ungauged (observation free) catchments. However, such relationships rely on the initial (gauged) model calibration, which is largely dependent on treatment of the various above mentioned error sources. As such, regionalization relationships also vary with how different errors are treated when establishing these relationships. In addition, because snow stations are selected primarily on the basis of the criteria of accessibility, safety of surveyors and maintainers, and protection from public disturbance [Molotch and Bales, 2005], the current snow observational network across the United States is sparse, leaving most snow-covered regions ungauged. For instance, there are over 1700 sites in the western United States at which SWE is measured. However, they are

still not sufficient to resolve SWE variability at the watershed scale [Bales *et al.*, 2006]. A comprehensive assessment of the impact of various uncertainty sources on snow model parameter behavior and regionalization relationships (among watershed characteristics and related snow model parameters) holds considerable potential for improved snowmelt predictions in ungauged snow-covered regions.

[5] The primary goal of the current study is to assess the impact of various sources of uncertainty on parameter behavior and model predictive uncertainty for a commonly applied snow model (SNOW17 of the National Weather Service (NWS)). We consider four different scenarios that differ in their underlying assumptions about the treatment of uncertainty. The secondary goal is to evaluate the impact of these error sources on the regionalization information of parameters. We utilize the recently developed DREAM algorithm [Vrugt *et al.*, 2008a, 2008b] and use different likelihood functions to define the four uncertainty scenarios. The DREAM algorithm is an adaptation of the SCEM-UA [Vrugt *et al.*, 2003] and has advantages over the SCEM-UA in the context of maintaining detailed balance and ergodicity [Vrugt *et al.*, 2008b, 2009] while also being efficient [Vrugt *et al.*, 2008b]. We compare posterior parameters for each likelihood function at 24 contrasting study sites. We are especially concerned with analyzing the uncertainty, correlation, and probability density of these posterior parameters. In addition, we evaluate model performance configured with different uncertainty sources in the context of providing SWE estimates at these climatically distinct sites. Furthermore, we develop a simple multiple linear regression model to construct potential regionalization relationships between SNOW17 parameters and site hydroclimatic characteristics under the various uncertainty scenarios considered.

2. Snow Model and Data Sets

[6] We apply the SNOW17 model at a range of Snow Telemetry (SNOTEL) sites maintained by the Natural Resources Conservation Service. It is generally recognized that a thorough investigation at the point scale provides critical insight on model behavior necessary for applying the model at larger scales [Hogue *et al.*, 2006a; Slater and Clark, 2006; Franz *et al.*, 2008; He, 2010]. In addition, the majority of ground-based snow observations utilized in operational forecasting are at the point scale. Although there are errors in SNOTEL observations, they are still the most widely used and longest record of snow information available across an extensive range of climatic and hydrologic regimes [Serreze *et al.*, 1999].

2.1. SNOW17 Model

[7] The NWS uses the SNOW17 model as part of their streamflow forecasting system in snow-dominated regions. The model uses empirically based relationships to simulate the heat storage of the snowpack, snowmelt, liquid water retention, and transmission. It requires air temperature and precipitation as inputs, while outputs include a rain plus snowmelt time series as well as SWE. SNOW17 is primarily controlled by 10 parameters when applied at the point scale (Figure 1 and Table 1) [Anderson, 2002; Franz *et al.*, 2008; He *et al.*, 2011].

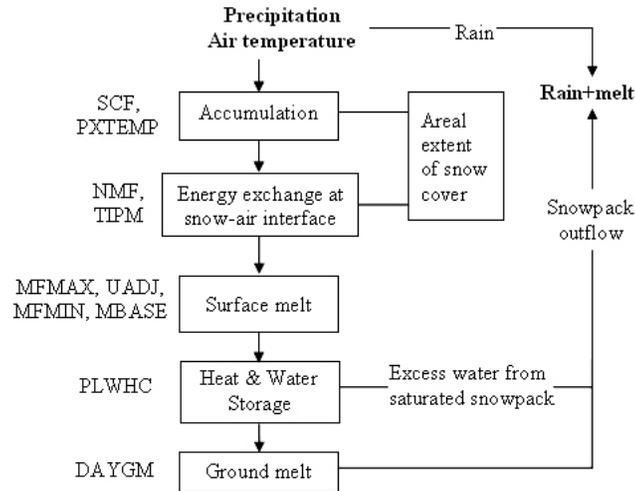


Figure 1. Schematic of SNOW17 model processes and corresponding parameters. The boxes designate model processes, and model inputs and output are highlighted in bold. Adapted from *Anderson* [1973].

[8] Gauge catch errors are accounted for using a snow correction factor (SCF, dimensionless). Snowfall input to the model, P_s (mm), is calculated as

$$P_s = P f_s \text{SCF}, \quad (1)$$

where P (mm) is the observed precipitation and f_s (dimensionless) represents the fraction of precipitation in the form of snow and is formulated as

$$f_s = \begin{cases} 1, & T_a \leq \text{PXTEMP} \\ 0, & T_a > \text{PXTEMP} \end{cases}, \quad (2)$$

where T_a ($^{\circ}\text{C}$) represents the observed air temperature and PXTEMP ($^{\circ}\text{C}$) is the temperature that distinguishes snowfall from rainfall. SCF (dimensionless) and PXTEMP ($^{\circ}\text{C}$) are the primary factors controlling snow input to the model. Heat exchange during nonmelt periods is controlled by parameters NMF (mm per 6 h per $^{\circ}\text{C}$) and TIPM (dimensionless) [Anderson, 1973]. Snowmelt during nonrain periods, M (mm), is calculated as

$$M = M_f(T_a - \text{MBASE}), \quad (3)$$

where M_f is the melt factor ($\text{mm } ^{\circ}\text{C}^{-1}$) and MBASE ($^{\circ}\text{C}$) signifies the threshold temperature above which snowmelt occurs. The melt factor is estimated through the use of a

sinusoidal function bounded by parameters MFMIN (mm per 6 h per $^{\circ}\text{C}$) and MFMAX (mm per 6 h per $^{\circ}\text{C}$). The parameter MFMIN dominates melt prior to 21 March, and MFMAX influences melt more significantly after 21 March [NWS, 2004]. During rain-on-snow periods, melt is primarily a function of the parameter UADJ (mm per mbar per $^{\circ}\text{C}$) which influences wind advection. The parameter DAYGM (mm d^{-1}) is used to characterize geothermal heat flux at the ground surface and to allow for a constant melt rate at the soil-snow interface. A more detailed description of the SNOW17 model is given by *Anderson* [1973] and *He et al.* [2011].

2.2. Study Sites

[9] The western United States contains an extensive SNOTEL network, which records measurements of precipitation, air temperature, SWE, and snow depth [Crook, 1977]. Despite the extensive spatial coverage, there have been only a few attempts to use SNOTEL data in research studies or operational forecasting [Serreze et al., 1999, 2001; Fassnacht et al., 2003; Franz, 2006; Slater and Clark, 2006] because of a range of concerns [Doesken and Schaefer, 1987; Fassnacht et al., 2003]. Serreze et al. [1999] developed a method to evaluate the quality of SNOTEL data, screen outliers, and eliminate negative precipitation and SWE data (also applied by Serreze et al. [2001] and Fassnacht et al. [2003]). In the current study, we utilize data from 24 SNOTEL sites (Table 2), focusing on

Table 1. Parameters of the SNOW17 Model With Ranges Estimated From *Anderson* [1973] and *Franz et al.* [2008]

Parameters	Explanation	Unit	Ranges
SCF	Snow fall correction factor	-	0.7–1.4
MFMAX	Maximum melt factor	mm per 6 h per $^{\circ}\text{C}$	0.5–2.0
MFMIN	Minimum melt factor	mm per 6 h per $^{\circ}\text{C}$	0.05–0.49
UADJ	The average wind function during rain-on-snow periods	mm per mbar per $^{\circ}\text{C}$	0.03–0.19
NMF	Maximum negative melt factor	mm per 6 h per $^{\circ}\text{C}$	0.05–0.50
MBASE	Base temperature for nonrain melt factor	$^{\circ}\text{C}$	0–1.0
PXTEMP	Temperature that separates rain from snow	$^{\circ}\text{C}$	–2.0–2.0
PLWHC	Percent of liquid water capacity	-	0.02–0.3
DAYGM	Daily melt at snow-soil interface	mm d^{-1}	0–0.3
TIPM	Antecedent snow temperature index parameter	-	0.1–1.0

Table 2. Summary of Selected SNOTEL Sites and Related Geophysical Information^a

Site	Site Name	State	Elevation (m)	Climate Region	Data Period (Water Year)	Maximum SWE (mm)	Annual Temperature (°C)	Annual Precipitation (mm)	Snowfall Fraction	Date of Maximum SWE	Date of Disappearance of Snow	Days of Snowmelt
MC	Mount Craig	Washington	1207	1	1991–2008	2073	4.35	1922	0.42	3 Apr	6 Jun	63
HM	Holland Meadows	Oregon	1503	1	1984–2008	1255	6.36	1964	0.34	3 Mar	18 May	66
LH	Lost Horse	Washington	1560	1	1991–2008	950	4.2	877	0.6	25 Mar	15 May	51
WC	Ward Creek #3	California	2028	2	1991–2008	1415	5.17	1845	0.58	2 Apr	27 May	55
IL	Independence Lake	California	2546	2	1995–2008	2068	4.38	2151 ^b	0.6	29 Apr	2 Jul	63
LL	Leavitt Lake	California	2931	2	1990–2008	2807	2.74	2503 ^b	0.64	25 Apr	4 Jul	69
ES	Emigrant Springs	Oregon	1158	3	1985–2008	312	7.65	865	0.2	12 Feb	11 Apr	58
TP	Tipon	Oregon	1570	3	1990–2008	439	5.37	589	0.52	18 Mar	2 May	44
MH	Mount Howard	Oregon	2411	3	1990–2008	927	1.73	1128	0.48	27 Apr	31 May	43
BK	Bisson Creek	Montana	1500	4	1993–2008	503	4.34	832	0.32	8 Apr	13 May	35
SM	South Mountain	Idaho	1981	4	1985–2008	828	6.44	919 ^b	0.52	18 Mar	6 May	48
ML	Meadow Lake	Idaho	2789	4	1990–2008	777	1.04	801	0.61	28 Apr	10 Jun	43
BC	Base Camp	Wyoming	2143	5	1989–2008	777	2.37	864	0.55	31 Mar	16 May	47
TD	Thumb Divide	Wyoming	2432	5	1990–2008	836	1.02	758	0.58	11 Apr	24 May	42
CB	Carrot Basin	Montana	2743	5	1984–2008	1194	0.36	1121	0.73	5 May	26 Jun	52
RC	Rock Creek	Utah	2404	6	1987–2008	452	4.1	585	0.44	26 Mar	26 Apr	31
LM	Lasal Mountain	Utah	2914	6	1989–2008	582	3.87	819	0.4	25 Mar	11 May	46
CP	Chepeta	Utah	3228	6	1990–2008	922	-0.14	792	0.54	15 Apr	27 May	43
CS	Crosho	Colorado	2774	7	1987–2008	516	3.06	717	0.47	5 Apr	10 May	35
BL	Brumley	Colorado	3231	7	1987–2008	485	-0.45	624	0.48	16 Apr	19 May	32
VL	Vallecito	Colorado	3316	7	1987–2008	874	2.56	865	0.57	5 Apr	20 May	44
WH	White Horse Lake	Arizona	2118	8	1990–2008	343	7.94	628	0.22	27 Feb	2 Apr	33
LN	Lookout Mountain	New Mexico	2591	8	1986–2008	104	8.68	450	0.15	7 Feb	2 Apr	54
SC	Silver Creek Divide	New Mexico	2743	8	1990–2008	630	5.64	671	0.44	4 Mar	19 Apr	45

^aStatistics are computed by water year (WY) and include average maximum snow water equivalent (SWE), date of maximum SWE, average date of disappearance of snow, average number of days of snowmelt (period from the day SWE peaks until complete melt out), and mean annual temperature and precipitation.

^bCorrected values obtained by applying the method of Serreze et al. [1999].

three elevation levels (low, median, and high) in each of eight predefined climatic regions [Serreze *et al.*, 1999] (Figure 2). Additionally, the 24 sites reside within, or in close proximity to, NWS forecast basins. Elevation of the selected study sites varies from 1158 m (site ES, Oregon) to 3316 m (site VL, Colorado) (Table 2).

2.3. Data Sets

[10] Quality control procedures outlined by Serreze *et al.* [1999] are used to screen SWE, precipitation, and temperature data from the 24 sites. Three of the sites (SM (Idaho), IL (California), and LL (California)) contain snowfall fractions (ratio of annual maximum SWE to annual total precipitation) greater than 1.0 for several years and are corrected using precipitation undercatch methods reported in the literature [Serreze *et al.*, 1999; He *et al.*, 2011]. In addition, about 1% of all data requires linear interpolation to fill short periods with missing data. After data processing, adequate time series are produced for periods ranging from 14 to 25 years. The final time period selected for study across all sites is water year (WY) 1995 to WY 2008 (14 years).

[11] Mean annual precipitation and maximum SWE at sites MC (Washington), HM (Oregon), WC (California), IL (California), and LL (California) are significantly higher than the other study sites (Table 2). These five sites are located in the windward side of the western mountains (Pacific Northwest and Sierra Nevada), where orographic effects deposit moist Pacific air [Mock, 1996]. These sites have an average melt period of 65 days. In comparison, the average melt period for the remaining sites in the interior regions is 44 days. Snowfall fractions at sites SC (New Mexico), LN (New Mexico), and WH (Arizona) (Arizona and New Mexico region) are generally lower, indicative of higher temperature and lower winter precipitation in this region. The SWE peak occurs earliest at site LN (New Mexico), while the peak occurs latest at site CB (Montana).

The snow melt-out date occurs earliest at sites LN (New Mexico) and WH (Arizona) and latest at sites IL (California) and LL (California). On average, most precipitation occurs in late fall and winter (November to March) at the study sites (Figure 3a). In contrast, precipitation in summer months (June to September) is much less. Air temperature at all sites shows significant seasonal variation (Figure 3b). In general, the snow accumulation season is from October to April, and the ablation season is from April to July (Figure 3c). These observations are generally consistent with the regional analysis of Serreze *et al.* [1999] for the same climate zones.

[12] The mean annual precipitation of the study sites ranges from 450 mm (site LN, New Mexico) to 2503 mm (site LL, California); mean annual temperature varies from -0.45°C (site BL, Colorado) to 8.68°C (site LN, New Mexico); mean annual snowfall fraction ranges from 0.15 (site LN, New Mexico) to 0.73 (site CB, Montana) (Table 2). Monthly precipitation, temperature, and SWE also show significant variations at different sites (Figure 3). The variability in our data sets exemplifies the fact that the selected sites cover diverse topographic and meteorological characteristics in the western United States and are spatially representative for this study.

3. Methods

3.1. SNOW17 Model Representation

[13] Let \tilde{X} and \tilde{Y} denote observed model forcing (i.e., precipitation and air temperature) and observed model output (i.e., SWE), respectively. Let \hat{X} denote model inputs (i.e., adjusted snowfall and air temperature) and let θ be the SNOW17 model parameters (the parameters tabulated in Table 1 except for SCF and PXTMP). The SNOW17 model can be written schematically as follows

$$\hat{Y} = f(\hat{X}, \theta). \quad (4)$$

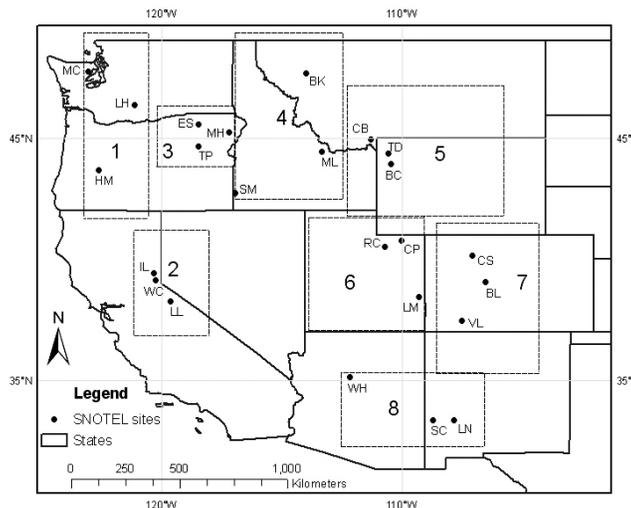


Figure 2. Location of 24 study sites (dots) as well as eight regions (boxes) that encompass the study sites. The regions and the median elevation of SNOTEL sites within the regions are 1, Pacific Northwest (1422 m); 2, Sierra Nevada (2439 m); 3, Blue Mountains, Oregon (1646 m); 4, Idaho and western Montana (1905 m); 5, NW Wyoming (2479 m); 6, Utah (2774 m); 7, Colorado (3037 m); 8, Arizona and New Mexico (2418m). Detailed information on the regions is given by Serreze *et al.* [1999].

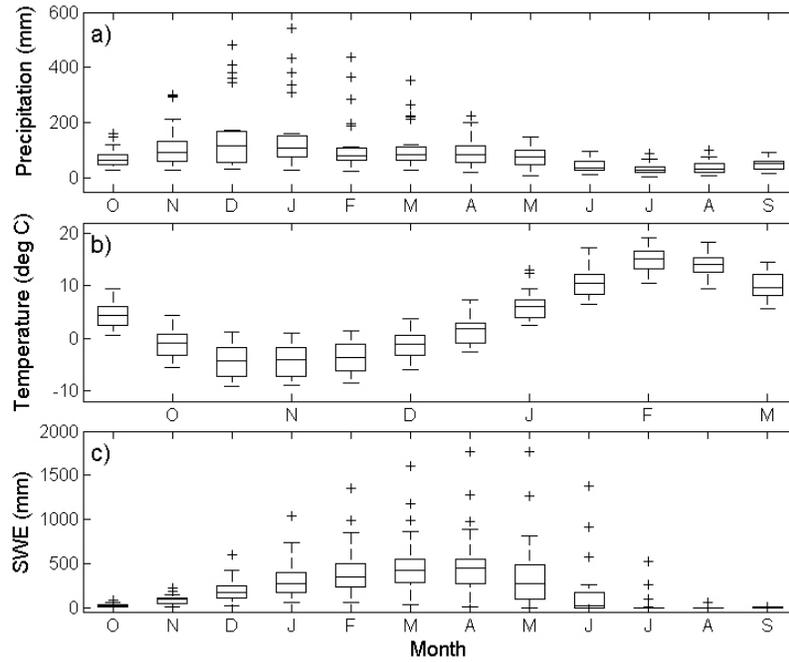


Figure 3. (a) Monthly precipitation, (b) temperature, and (c) snow water equivalent (SWE) for the 24 SNOTEL sites presented in Table 2.

[14] It is worth noting the following: first, \hat{X} is computed from \tilde{X} via SCF and PXTEMP through the relationship stated in equations (1) and (2); second, all simulations in this study start at a zero SWE (i.e., October 1), therefore, all SNOW17 states are zero at the beginning of the simulation. Initial conditions are thus not included in equation (4).

[15] To assess if f accurately represents measured snow dynamics, it is a standard practice to compare measured \tilde{Y} and simulated \hat{Y} via the following error vector:

$$\mathbf{e}(\theta) = \{\varepsilon_1(\theta), \varepsilon_2(\theta), \dots, \varepsilon_T(\theta)\}, \quad (5)$$

where $\varepsilon_t(\theta)$ designates the residual at day t , $\varepsilon_t(\theta) = \hat{y}_t - \tilde{y}_t$, $t = 1, 2, \dots, T$, and T represents the total number of days of the simulation period. For convenience, this vector is typically aggregated in a single error term, the sum of squared residuals (SSR)

$$\text{SSR} = \sum_{t=1}^T \varepsilon_t(\theta)^2. \quad (6)$$

[16] Minimization of SSR is typically achieved by tuning (calibrating) model parameters, without consideration of errors in model forcing, calibration data, and model structure (i.e., inadequacies in f).

3.2. Bayesian Inference of Posterior Probability Density Function of Model Parameters

[17] In the past decade, various optimization algorithms have been developed to minimize equation (6). These algorithms either provide an estimate of the optimal parameter set (e.g., SCE-UA of Duan *et al.* [1992]) or derive the entire underlying posterior probability density function (pdf) of

the parameters (e.g., SCEM-UA of Vrugt *et al.* [2003]). The posterior pdf helps simulate predictive uncertainty of f . However, it would be more ideal to evaluate the posterior pdf of θ and the predictive distribution of \tilde{Y} in the presence of other uncertainty sources (i.e., data forcing, model structure, and SWE measurements). This can be achieved in the context of Bayesian statistics coupled with Monte Carlo sampling. The Bayesian paradigm offers the flexibility of combining multiple sources of uncertainty on the basis of Bayes' theorem. Following the standard Bayesian law, the posterior pdf of the SNOW17 parameters after conditioning on the observed SWE (\tilde{Y}), $p(\theta|\tilde{Y})$, can be expressed as

$$p(\theta|\tilde{Y}) = \frac{p(\tilde{Y}|\theta)p(\theta)}{p(\tilde{Y})}, \quad (7)$$

where $p(\tilde{Y}|\theta)$, $p(\theta)$, and $p(\tilde{Y})$ represent the data likelihood, a prior parameter distribution, and the probability density of observing the SWE, respectively. It is usually infeasible to directly calculate $p(\tilde{Y})$. The typical practice is to reformulate equation (7) as follows:

$$p(\theta|\tilde{Y}) \propto L(\theta|\tilde{Y})p(\theta), \quad (8)$$

where $L(\theta|\tilde{Y}) \equiv p(\tilde{Y}|\theta)$ is the likelihood function.

3.3. Assessment of Model Forcing, Structure, and Measurement Uncertainty

3.3.1. Forcing Uncertainty

[18] Precipitation and air temperature are the primary forcing variables controlling snow dynamics. Consequently, these variables exert primary control on the simulated values of SWE for a given model [Durand and

Margulis, 2006, 2008]. In recent years a few papers have appeared in the literature focusing on quantification of precipitation uncertainty. A common approach is to assume that precipitation is corrupted by lognormally distributed multiplicative errors and then estimate the multiplier [e.g., *Margulis et al.*, 2002; *Kavetski et al.*, 2006; *Ajami et al.*, 2007]. However, SNOW17 already applies a snow correction factor to adjust precipitation input (i.e., equation (1)). Assigning another multiplier to the precipitation forcing for the model would be redundant. An alternative approach is to apply a single multiplier to each individual precipitation event and estimate these multipliers along with the model parameters [e.g., *Kavetski et al.*, 2002; *Vrugt et al.*, 2008b]. This method, however, is computationally expensive, especially for CPU-intensive forward models with numerous precipitation events. Geostatistical methods are also reported in the literature in the context of estimating precipitation, with uncertainty accounted for via conditional simulation [*Clark and Slater*, 2006; *Gotzinger and Bardossy*, 2008]. However, these methods are not straightforward to implement in practice and thus are not widely applied.

[19] The current study adopts an alternative, but simpler, approach to address precipitation uncertainty associated with SNOW17 simulations. This approach draws inspiration from the multiplier approach discussed above and uses the parameters SCF and PXTEMP to quantify snowfall uncertainty. The posterior distribution of these two parameters can be obtained by comparing measured and simulated SWE, as discussed in section 3.3.2.

3.3.2. Measurement Uncertainty

[20] Measurement uncertainty consists of system error (resulting from the instruments) and representation error (resulting from aggregation of spatial variability in measured variables). Since this study focuses on the point scale, the impact of aggregation error can be neglected. We also assume mutually independent and normally distributed errors of the SWE observations.

[21] We apply a Bayesian approach to estimating parameter uncertainty in the presence of measurement uncertainty, forcing data error, and model structural errors. In the first trial, we use a standard log likelihood function to estimate the posterior distribution of the SNOW17 parameters using the SWE measurements and a uniform prior distribution:

$$\lambda(\theta|\tilde{Y}) = -\frac{T}{2}\ln(2\pi) - \frac{T}{2}\ln(\sigma^2) - \frac{1}{2\sigma^2}\sum_{t=1}^T[\varepsilon_t(\theta)]^2, \quad (9)$$

where $\lambda(\theta|\tilde{Y})$ is the log likelihood function and σ denotes the error standard deviation of the SWE measurements. Equation (9) has been successfully applied in different fields of study, including rainfall-runoff modeling [*Vrugt et al.*, 2008a] and hydrogeophysics [*Huisman et al.*, 2010].

3.3.3. Measurement and Model Structural Uncertainty

[22] Despite its wide application, the assumption of mutually independent errors is not very realistic since the SWE error residuals in equation (5) generally exhibit significant autocorrelation. This serial dependence of the errors needs to be explicitly considered during model calibration if our goal is to derive meaningful probability distributions of the SNOW17 parameters and maximize chances of finding useful regionalization relationships. A

relatively simple and effective method to account for auto-correlated error residuals is the use of a first-order autoregressive model (AR(1)) of the error residuals:

$$\varepsilon_t(\theta) = \rho\varepsilon_{t-1}(\theta) + v, \quad t = 1, 2, \dots, T, \quad (10)$$

where ρ denotes the first-order correlation coefficient and v is the remaining (unexplained) error with zero mean and constant standard deviation σ_v . We follow *Sorooshian and Dracup* [1980] and directly incorporate the AR(1) model into the log likelihood function (equation (9))

$$\begin{aligned} \lambda(\theta, \rho, \sigma_v|\tilde{Y}) = & -\frac{T}{2}\ln(2\pi) - \frac{1}{2}\ln\left(\frac{\sigma_v^2}{1-\rho^2}\right) \\ & - \frac{T}{2}(1-\rho^2)\sigma_v^{-2}\varepsilon_1(\theta)^2 - \frac{1}{2}\sigma_v^{-2}\sum_{t=2}^T\delta_t(\theta, \rho, \sigma_v)^2, \end{aligned} \quad (11)$$

where

$$\delta_t(\theta, \rho, \sigma_v) = \varepsilon_t(\theta) - \rho\varepsilon_{t-1}(\theta), \quad t = 1, 2, \dots, T, \quad (12)$$

is the AR(1) time series of residuals of SWE with $\varepsilon_0 = 0$. This approach at least partially accounts for the effect of model error, and has been successfully applied in watershed modeling [*Vrugt et al.*, 2008a]. The variables ρ and σ_v of the AR(1) model are additional parameters to be estimated along with SNOW17 model parameters. Augmenting θ in equation (11) to include both SCF and PXTEMP further accounts for errors in snowfall (precipitation) forcing.

3.3.4. Uncertainty Scenarios

[23] Four different uncertainty scenarios (S1–S4), with increasing complexity, are developed in this study. S1 only considers uncertainty in the SNOW17 parameters, S2 jointly considers parameter and SWE measurement uncertainty, S3 is similar to S2 but also includes explicit treatment of model forcing data error, and S4 is the most comprehensive in that it considers model structural errors in addition to model parameters, SWE measurement, and the forcing data uncertainty addressed previously in S3. In fact, S1 is similar to the traditional calibration approach, which attributes all potential sources of modeling error to parameter uncertainty. S4 is the more ideal (and realistic) case and attempts to disentangle the main sources of error involved in hydrologic modeling. Equation (9) is used as the likelihood function for both S2 and S3. Equations (6) and (11) are utilized as the log likelihood function for S1 and S4, respectively. A comprehensive summary of these four scenarios is given in Table 3.

Table 3. Uncertainty Scenarios Considered^a

Scenarios	Uncertainty Sources Considered	Likelihood Function	Number of Parameters Considered
S1	U1	Equation (6)	8
S2	U1 + U2	Equation (9)	9
S3	U1 + U2 + U3	Equation (9)	11
S4	U1 + U2 + U3 + U4	Equation (11)	12

^aU1 stands for SNOW17 parameters; U2 designates SWE measurements; U3 represents precipitation forcing; U4 denotes model structure.

3.4. Implementation of DREAM in SNOW17 Modeling

[24] In the absence of closed form analytical solutions of the posterior pdf of the parameters, we use Markov chain Monte Carlo simulation with DREAM to summarize model parameters and predictive uncertainty. The DREAM runs N (predefined) different Markov chains in parallel. If the state of a single chain is given by a single d -dimensional vector, then the N chains define a population Θ , which corresponds to an $N \times d$ matrix. Jumps in each chain are generated using a discrete proposal distribution that computes a fixed multiple of the difference of the states of randomly chosen pairs of other chains. This d -dimensional difference vector contains the desired information about the scale and orientation of the proposal distribution. By accepting each jump with the Metropolis ratio, a Markov chain is obtained whose stationary distribution is the posterior target distribution. The proof of this is given by *Vrugt et al.* [2008a]. The convergence of a DREAM run is monitored with the \hat{R} statistic of *Gelman and Rubin* [1992], which compares the variance within and between the chains ($\hat{R}_k < 1.2$, $k = 1, 2, \dots, d$, designates convergence to a limiting distribution). The reader is referred to *Vrugt et al.* [2008a] for a detailed description of DREAM. Using different case studies, *Vrugt et al.* [2009] showed that DREAM achieves good sampling efficiencies for a large range of problems.

[25] In this study, DREAM is applied to estimate the posterior distribution of the SNOW17 parameters using standard settings of the algorithmic variables. Equations (6), (9), and (11) are applied as the (log) likelihood functions for uncertainty scenarios S1, S2 and S3, and S4, respectively (Table 3). The numbers of parameters considered in the four different modeling scenarios (S1–S4) are 8 for S1, 9 (S1 with σ) for S2, 11 (S2 with SCF and PXTMP) for S3, and 12 (S3 with σ_v) for S4 (Table 3). For each DREAM simulation, we use a maximum total of 100,000 model evaluations. Preliminary tests have shown that this number is sufficient to obtain convergence to a stationary distribution according to the \hat{R} statistic.

[26] The uncertainty analysis is presented in terms of (1) 95% parameter uncertainty, (2) marginal distribution and correlation of posterior parameters, and (3) 95% SWE prediction uncertainty under each uncertainty scenario. In assessing the 95% parameter uncertainty, the focus is on parameter uncertainty width and the correlation between the width and various hydroclimatic characteristics of the different study sites. Traditional linear correlation coefficients (which range from -1 to 1 , with a higher absolute value indicating a stronger correlation) are used to quantify the strength of the relationship. A corresponding p value statistic is also estimated to determine if the correlation is statistically significant or not. A value of $p < 0.05$ indicates the presence of significant correlation.

3.5. Multiple Linear Regression Model

[27] Multiple linear regression (MLR) has found widespread application in modeling the relationship between a dependent variable and multiple independent variables. Recent applications of the MLR include regionalization of hydrologic model parameters [*Campbell and Bates*, 2001; *Heuvelmans et al.*, 2006; *Bastola et al.*, 2008], water level and streamflow forecasting [*Seidou and Ouarda*, 2007;

Tiwari and Chatterjee, 2010], water temperature forecasting [*Salomon and Feyen*, 2009], and flood frequency analysis at ungauged sites [*Leclerc and Ouarda*, 2007]. In this study we apply the MLR to establish regionalization relationships and assess the dependence between the DREAM-derived maximum likelihood values of the SNOW17 parameters (which exhibit a maximum value of the log likelihood function) and respective SNOTEL site characteristics

$$\theta = A \times S + C, \quad (13)$$

where $\theta = [\theta_1, \theta_2, \dots, \theta_d]$ represents the dependent SNOW17 model parameters (Table 1), $S = [s_1, s_2, \dots, s_m]$ designates the independent site characteristics considered (Table 2), m is the total number of characteristics ($m = 6$ in this study), A is a regression coefficient matrix with a size of $d \times m$, and $C = [c_1, c_2, \dots, c_m]$ is a vector with constant values. MLR is a rather simplistic modeling approach and is restricted to linear dependencies between input and output variables. However, the main advantage is that the results are easy to interpret and implement. On the contrary, curve-fitting methods such as artificial neural networks (ANNs) can provide (nonlinear) mapping between variables, but their (black box) results are more difficult to generalize and interpret. Furthermore, despite its conceptual simplicity, the resulting error of MLR is often comparatively small compared to more complicated techniques such as ANNs [*Heuvelmans et al.*, 2006; *Bastola et al.*, 2008]. We therefore restrict our current analysis to MLR and will consider ANNs in future work.

4. Results and Discussion

4.1. Uncertainty of Posterior Parameters

[28] Uncertainty of the posterior parameters is investigated in two ways. First, the upper bound and lower bound of each parameter derived from DREAM under the four scenarios are assessed. The assessment specifically focuses on 95% parameter uncertainty and aims to evaluate the impacts of various error sources in snow modeling on SNOW17 parameter behavior. Second, the correlation between various site characteristics and posterior parameter ensemble width (uncertainty) is evaluated. The overall goal of this evaluation is to identify potential regional relationships between site hydroclimatic conditions and parameter uncertainty.

[29] When all uncertainty sources are considered (S4), 95% parameter uncertainty tends to cover most of the feasible parameter range for each parameter at three selected sites (Figure 4). These three sites, WH (Arizona) from the Arizona–New Mexico region, BL from Colorado, and LL (California) from the Pacific Northwest region, represent a range of contrasting annual snowfall behavior, including scarce, medium, and high snowfall, respectively (Table 2). In comparison, under the other three scenarios (S1, S2, and S3), the 95% parameter uncertainty intervals are significantly narrower for all parameters. This contrast illustrates that model structural uncertainty considerably impacts parameter uncertainty and dominates the total uncertainty involved in SNOW17 modeling. Despite the narrow uncertainty interval, the 95% parameter uncertainty derived from S3 differs from that derived from S1 and S2. This is evident in terms of both the location of the interval in the feasible

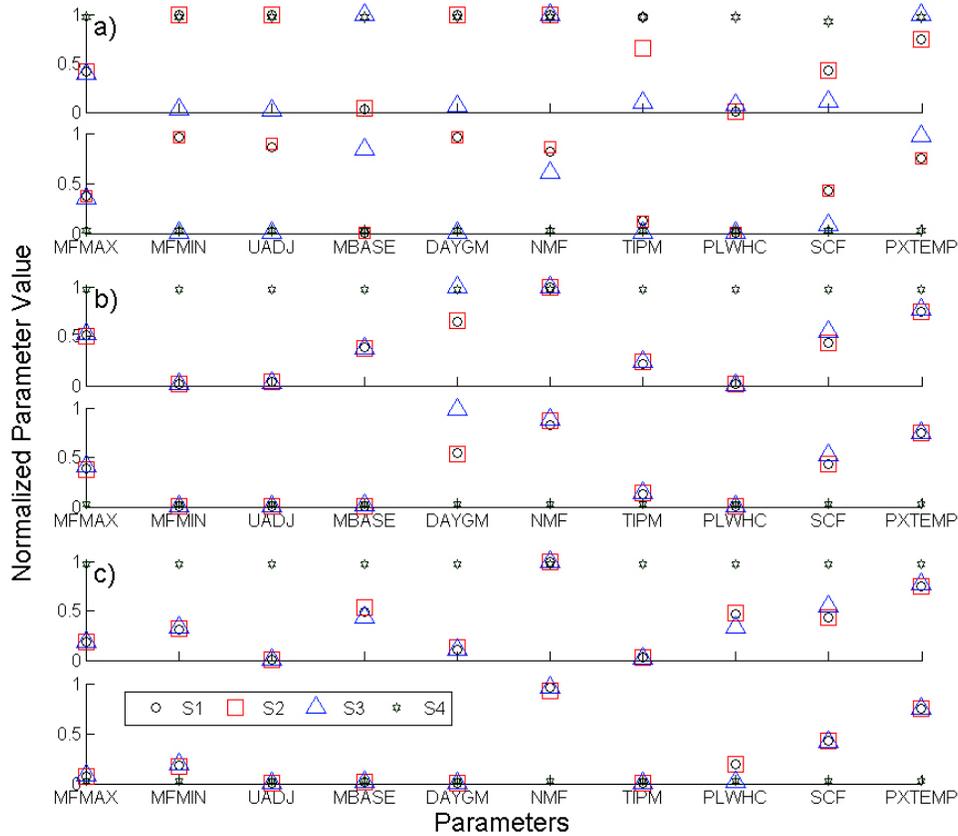


Figure 4. Normalized 95% posterior parameter range for sites (a) Leavitt Lake (LL, California), (b) Brumley (BL, Colorado), and (c) White Horse Lake (WH, Arizona). For each site, the top and bottom plots show the upper bound and lower bound of posterior parameters for the four uncertainty scenarios, respectively. The x axis designates SNOW17 parameters; the y axis denotes normalized parameter values (normalized by the feasible range of each parameter).

parameter range (e.g., parameters MFMIN, UADJ, MBASE, DAYGM, and NMF at site LL; Figure 4a) and the magnitude of the interval (e.g., parameter NMF at site LL; Figure 4a). This finding indicates that uncertainty in snowfall forcing is the secondary factor (after the primary structural uncertainty) impacting parameter uncertainty. This impact is most apparent at the wet site (Figure 4a) and gradually diminishes as the study sites become drier (Figures 4b and 4c). The 95% parameter uncertainty associated with scenarios S1 and S2 is almost identical, with only a few exceptions at the wet site (i.e., parameter TIPM at site LL; Figure 4a); hence, explicit consideration of Gaussian noise in the measurements (S2) has marginal impacts on SNOW17 parameter uncertainty.

[30] As mentioned, parameter uncertainty varies from site to site under each scenario. This suggests that the physical site characteristics (topographic and climatic factors) likely influence parameter uncertainty. To more closely examine this potential influence, correlations between posterior parameter uncertainty interval (width) and site information are evaluated (Figure 5). In general, there are only a few site-parameter combinations that show fairly significant correlation (absolute correlation value greater than 0.33) under each scenario. For S1 (Figure 5a), the uncertainty in PLWHC is negatively correlated to precipitation (−0.41), maximum SWE (−0.41), and melting period

(−0.37). Recall that PLWHC determines the liquid water capacity of the snowpack. When the capacity is met, snowmelt leaves the snowpack. More precipitation and higher maximum SWE are more likely to lead to higher snowmelt, given that the liquid water capacity of the snowpack can be more rapidly filled. PLWHC thus has less (not primary) impact on the determination of snowpack runoff generation, leading to smaller variations in its range. This further indicates that parameter PLWHC is more identifiable in wetter regions (more precipitation). The negative correlation between PLWHC uncertainty and melting period is most likely indirect, built up via the correlation between precipitation and melting period. For S2 (Figure 5b), negative correlation is also observed between PLWHC uncertainty and the same set of site characteristics as in S1, indicative of the marginal influence of SWE measurement uncertainty on the potential relationship between parameter uncertainty and site characteristics.

[31] PXTEMP uncertainty is negatively related to maximum SWE, precipitation, and the melting period under S3 (Figure 5c). Recall that PXTEMP (ranging from −2°C to 2°C) distinguishes snowfall from rainfall in precipitation (precipitation is 100% snowfall if air temperature is less than PXTEMP; otherwise, precipitation is 100% rainfall). A stable and high PXTEMP value (i.e., close to the upper bound of 2°C) would produce potentially higher maximum

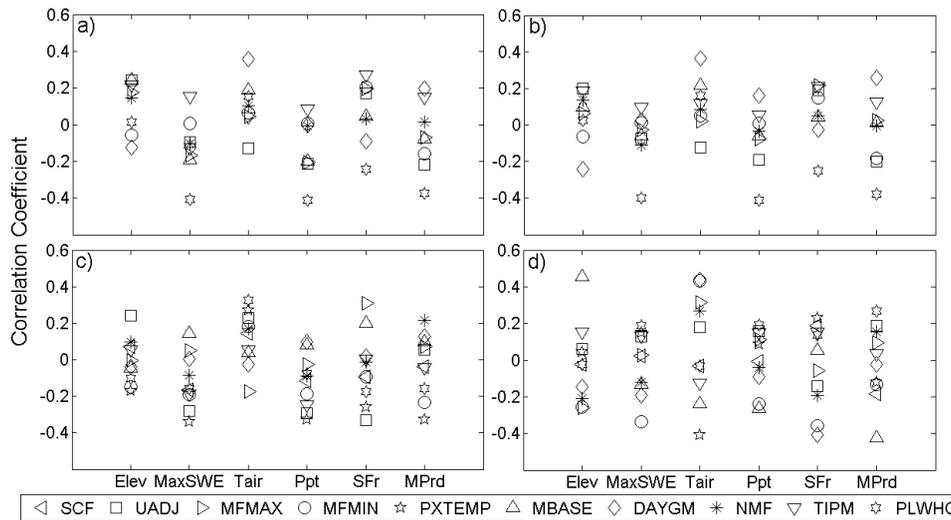


Figure 5. Correlation between the DREAM-derived posterior parameter ranges and selected topographic and climatic characteristics (x axis) associated with each study site for the SNOW17 model parameters under uncertainty scenarios (a) S1, (b) S2, (c) S3, and (d) S4. These site characteristics include elevation (Elev), maximum annual SWE (MaxSWE), mean annual air temperature (Tair) and precipitation (Ppt), snowfall fraction (SFr), and the length of melting period in days (MPrd). Parameters SCF and PXTEMP are fixed in both S1 and S2 and are thus not included in Figures 5a and 5b.

SWE values because more precipitation would fall as snow. In this case, PXTEMP exhibits only slight variations in its range. In comparison, more site-parameter uncertainty combinations show fairly significant correlation for S4 (Figure 5d). First, DAYGM uncertainty is positively correlated to air temperature (0.44) and negatively correlated with the snowfall fraction (-0.41). As mentioned, DAYGM is a constant daily rate of melt at the soil-snow interface. Higher air temperature leads to more heat exchange with the ground surface and thus more snowmelt at the snowpack-ground interface. Consequently, DAYGM tends to vary in a larger range, generating more melt. The correlation between DAYGM uncertainty and the snowfall fraction is likely constructed via the correlation between air temperature and the snowfall fraction (higher temperature results in smaller snowfall fraction of precipitation). Second, uncertainty in MBASE is positively correlated to elevation (0.45) and is negatively related to melting period (-0.43). MBASE is a temperature parameter above which nonrain melt will occur. A longer melting period is typically characterized by a more stable nonrain melting process. In this situation, the variation in MBASE would be expected to be low. Third, the uncertainty of PXTEMP is negatively related to air temperature (-0.41). A potential explanation for this is that the chance of snowfall decreases with increasing temperature and, hence, at higher temperatures PXTEMP will have less impact on the partitioning between snowfall and precipitation. Fourth, uncertainty in MFMIN is positively correlated to air temperature and negatively related to the snowfall fraction. Recall that MFMIN mostly controls the determination of nonrain melt rate before 21 March. Higher air temperature likely increases midwinter melt (before 21 March), causing larger variations of the value of MFMIN. The correlation between the uncertainty of MFMIN and the snowfall fraction is most likely indirect, constructed via air temperature.

[32] Of the four scenarios, S1 and S2 generally have very similar site-parameter uncertainty correlations (Figures 5a and 5b). This further illustrates that including Gaussian noise in SWE measurement (S2) has relatively minor impact on parameter uncertainty. However, when forcing uncertainty is explicitly considered (S3), the correlation relationships change significantly (Figure 5c), and overall correlations increase. When model structural errors are included (S4), the overall correlations increase, with more pairs of site characteristics and parameter uncertainty displaying fairly strong correlations (Figure 5d). However, it is not straightforward to characterize these changes in correlations caused by the inclusion of snowfall uncertainty or model structure uncertainty.

4.2. Distribution and Correlation of Posterior Parameters

[33] The normalized distributions of the posterior parameters at sites LL (California), BL (Colorado), and WH (Arizona) for the four different uncertainty scenarios are graphically depicted in Figure 6. Normalization with the prior range is utilized to facilitate comparison of the posterior parameter distributions of the four different scenarios. In general, three types of marginal distributions are observed: uniform, normal, and lognormal. It is also evident that parameter distributions at all three sites in S1 are very similar to those of S2, indicative of the marginal impact of Gaussian measurement uncertainty on the distribution of posterior parameters. In addition, posterior parameters in S3 at the three sites by and large have similar types of distributions as the parameters in S1 and S2. However, most of the parameters have different distribution characteristics (e.g., the skewness and moments). This suggests that including uncertainty in snowfall input tends to generate similar 95% posterior parameter ranges, but with different shapes of the respective marginal distributions.

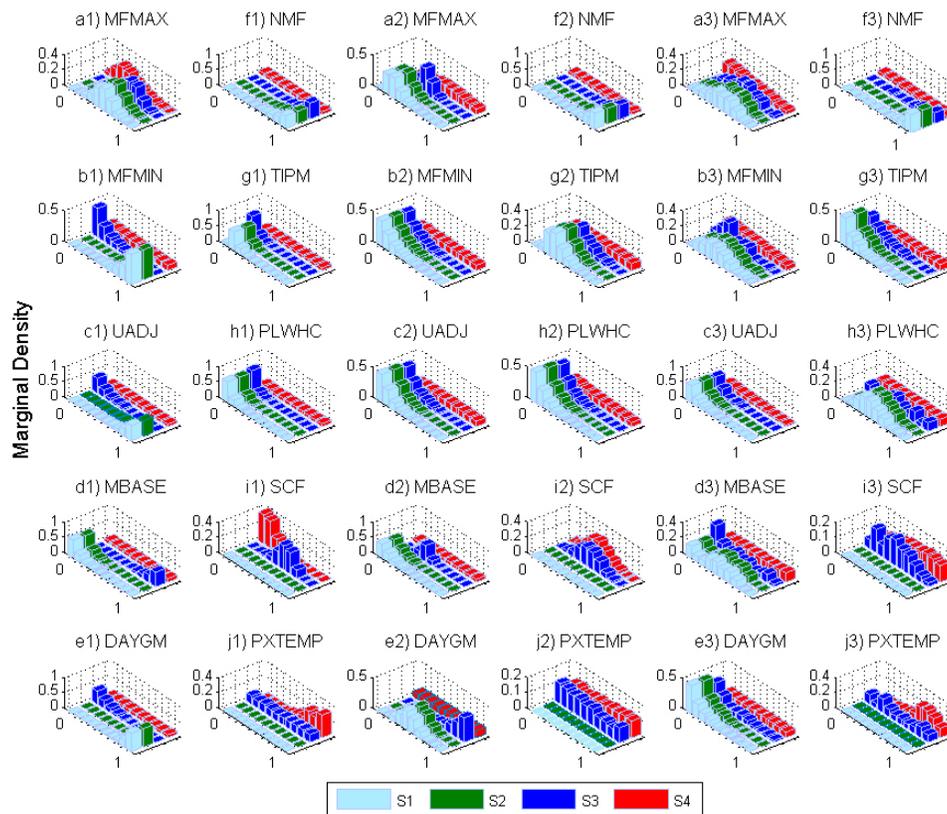


Figure 6. Normalized posterior marginal parameter distributions at three selected sites: Leavitt Lake (LL, California, distributions a1 to j1), Brumley (BL, Colorado, distributions a2 to j2), and White Horse Lake (WH, Arizona, distributions a3 to j3) under four uncertainty scenarios. The x axis denotes scenarios (S1 to S4 from left to right); the y axis designates normalized posterior parameter ranges (from 0 to 1); the z axis shows the marginal density of posterior parameters.

Furthermore, except for the parameters controlling snowfall input (i.e., SCF and PXTEMP) and a melt parameter (i.e., MFMAX), posterior parameters in S4 follow a uniform distribution at the three different sites. Model structural errors appear to be the main source of error, and parameter values can be found that extend the entire prior defined ranges that compensate for this structural deficiency.

[34] Parameters MFMAX, UADJ, and NMF generally follow the same types of distribution at the three sites under scenarios S1, S2, and S3 (Figure 6), while the distribution of MFMIN, MBASE, DAYGM, TIPM, and PLWHC varies among sites (normal or lognormal) under the three scenarios. Parameters SCF and PXTEMP roughly follow normal and uniform distributions at the three sites for S3, respectively. For a specific site, for example, site LL (California), different parameters have different types of distributions (Figure 6, distributions a1–j1). This highlights the complex form of the joint distribution of the posterior parameters.

[35] There are relatively few significant correlations among the posterior parameters under each uncertainty scenario across the 24 study sites (Figure 7). Overall, the percentage of parameter pairs with strong correlation (defined here as the absolute value of correlation coefficient greater than 0.67) is consistently less than or equal to 20% at all sites in all four scenarios. S1 and S2 have very similar correlation structure except for at site CS (Colorado). This may be due to higher uncertainty (lower accuracy) in the

SWE measurement at site CS. Introducing uncertainty in the measurement forces the model to overtune parameters and potentially alter interactions between parameters. In addition, there are consistently no significant correlations for both S1 and S2 at one third of the study sites (Figure 7). These sites include the three wettest sites (i.e., HM (Oregon), IL (California), and LL (California)) and another five sites with medium wetness. It is likely that SNOW17 parameters perform more independently in wet rather than extremely dry conditions. The percentage of significant correlations associated with S3 is apparently different from that of S1 and S2, suggesting that inclusion of snowfall uncertainty largely alters parameter interaction relationships. Particularly, at sites TD (Wyoming) and MC (Washington), 20% of the posterior parameter pairs show significant correlations. Under S4, the percentage of significant correlation is considerably less than S3 at most study sites. The highest percentage observed is 4.4%. This is expected since as illustrated in Figure 6, including model structural error in our analysis leads to much wider parameter ranges. The variance of the posterior parameter is thus larger, leading to lower correlations among those respective posterior parameters. The site mean percentages of significant correlations for the four scenarios are calculated as 3.87%, 4.91%, 7.78%, and 3.06%, respectively. This indicates that inclusion of uncertainty in SNOW17 model forcing (S3) tends to promote parameter correlation, followed

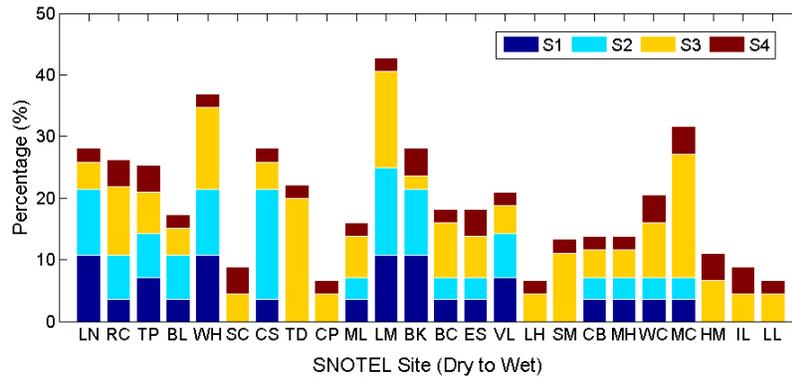


Figure 7. Percentage of posterior parameter pairs with strong correlation (the absolute correlation value is greater than 0.67) under the four uncertainty scenarios.

by consideration of SWE measurement uncertainty (S2), parameter uncertainty (S1), and model structural uncertainty (S4). Furthermore, for all four scenarios considered herein we could not find any apparent relationship between the percentage of significant correlation and site wetness (Figure 7). This signifies the difficulty in generalizing the dependence (if present at all) of posterior parameter inter-correlation on site climatic characteristics.

4.3. Prediction of Posterior Parameter Sets

[36] Performance of the posterior parameters under each uncertainty scenario is examined and demonstrated at three sites, LL (California), BL (Colorado), and WH (Arizona) (Figure 8). The focus is on 95% SWE prediction uncertainty

derived from the parameters in three years with differing climatology. The selected years correspond to a relatively wet year, a normal year, and a dry year with respect to snowfall received at each site. As expected, posterior parameters from S4 produce wide SWE predictions, which generally cover all the observed SWE at the three sites. Alternatively, S1, S2, and S3 have much narrower posterior parameter bounds (Figure 4) and produce much narrower ensembles of SWE. Furthermore, the SWE estimates from the four scenarios generally mimic the patterns of observed SWE. However, the model-predicted SWE bound associated with S3 more closely and more appropriately matches the variability in SWE measurements observed at the three different sites for the three selected years. In addition, for all three years at

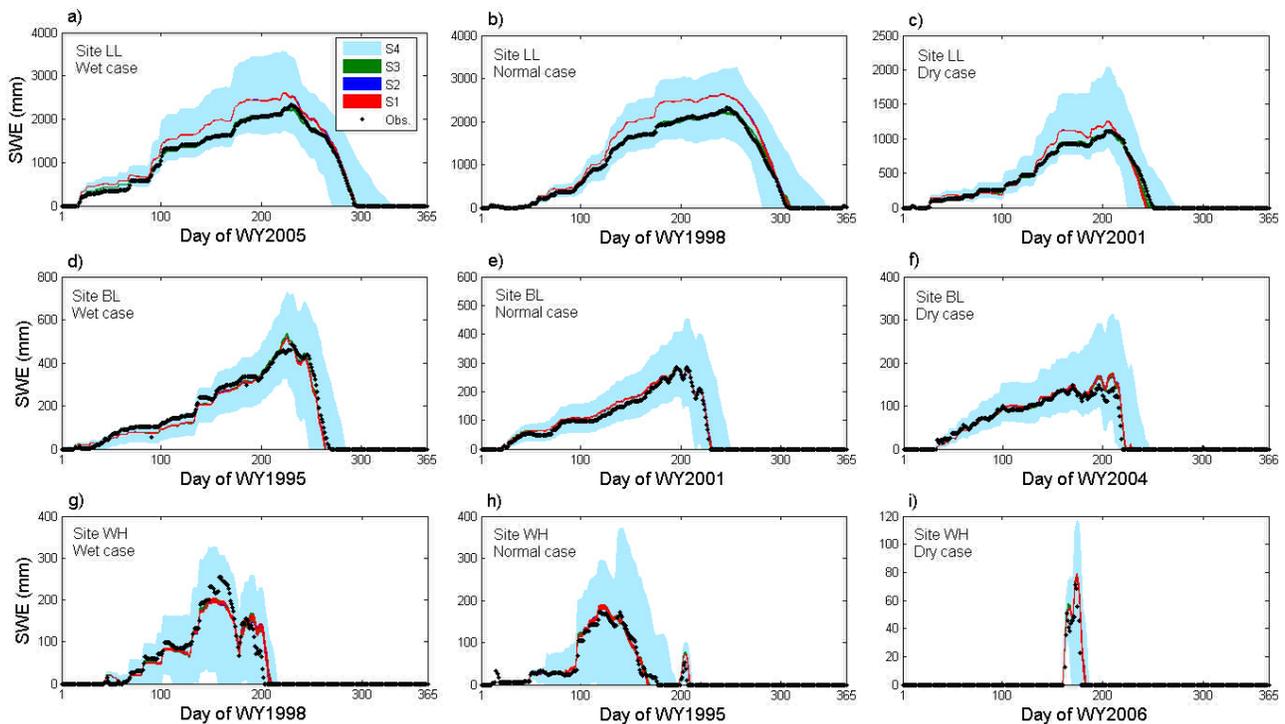


Figure 8. The 95% SWE prediction uncertainty ranges (shaded regions) along with SWE observations (dots). Results for (a–c) Leavitt Lake (LL, California), (d–f) Brumley (BL, Colorado), and (g–i) White Horse Lake (WH, Arizona). Examples of model performance for (left) wet, (middle) normal, and (right) dry years at each site are shown. Note the vertical scale differences at the three sites.

sites LL (California) and BL (Colorado), the melt-out dates of the SWE predictions from S4 are significantly lagged in comparison to the SWE observations. This is because the overestimated peak SWE predictions take much longer to melt out in S4.

[37] At site LL (California), the SWE prediction of S3 more accurately captures the peak and the overall variations of SWE measurement in all three years (Figures 8a–8c) when compared to the uncertainty bounds associated with S1 and S2. In comparison, at the two drier sites, BL (Colorado) and WH (Arizona), SWE predictions from S1, S2, and S3 are nearly identical (Figures 8d–8i). Apparently, the explicit consideration of uncertainty in snowfall input significantly improves SWE estimates at the wet site. However, at the drier site, this gain in skill is somewhat limited. It is worth noting that at site WH (Arizona), in the wet year, SWE ensemble estimates from S1, S2, and S3 miss the observed peak SWE. The ablation process is also overestimated. These results illustrate that it is possible to produce SWE ensemble estimates without satisfactory coverage of the observations. We posit that this is caused by structural errors in the model that become dominant when confronted with high temporal variability of snow accumulation and ablation.

4.4. Regionalization of SNOW17 Parameters

[38] The functional relationships between the DREAM-derived maximum likelihood parameters and SNOTEL site characteristics are determined using MLR (equation (13)). The correlations between parameters derived from these MLR relationships (regression-based) and maximum likelihood parameters under all four scenarios are calculated (Table 4). Overall, the correlations are rather weak, with a mean value of 0.54, 0.52, 0.50, and 0.51 for S1, S2, S3, and S4, respectively. The low correlations observed here are similar to the findings of previous studies that focused on a significantly large number (more than 300) of study sites [e.g., Merz and Bloschl, 2004; Parajka et al., 2005]. One potential explanation of these low correlations is that some of the site characteristics considered in the study (e.g., elevation and melting period) may not be directly and strongly relevant to snow accumulation and ablation processes. Nonetheless, there are several cases with significant correlation (Table 4). For instance, under S1 and S2, significant regionalization relationships are found for parameters

UADJ and DAYGM. For S4, strong correlations are also observed for MBASE and PXTEMP. Under S3, regression-based PXTEMP is significantly related to DREAM-derived PXTEMP (0.74; *p* value of 0.02).

[39] To closely examine these exceptions and investigate the impact of various uncertainty configurations on regionalization information of SNOW17 parameters, the MLR-derived relationships corresponding to those four parameters (UADJ, MBASE, DAYGM, and PXTEMP) under all uncertainty scenarios are presented (Figure 9). Overall, all four parameters are generally correlated to site air temperature and the snowfall fraction. This is not surprising, as parameters UADJ, MBASE, and DAYGM all contribute to the determination of snowmelt, which is a function of site air temperature and the amount of snowfall available, while PXTEMP is a temperature index parameter that controls the partitioning of snowfall from precipitation. Parameter MBASE is also shown to be related to the melting period. This correlation is most likely indirect and induced by the amount of snowfall. For those cases associated with significant correlations (with absolute correlation greater than 0.67; Table 4), the regionalization relationships of parameters UADJ (Figure 9a) and DAYGM (Figure 9c) are very similar under S1 and S2. Strong regionalization relationships are apparently insensitive to explicit inclusion of SWE measurement uncertainty (S2). Parameter MBASE is correlated to the snowfall fraction and the melting period (Figure 9b). Parameter PXTEMP is related to different site characteristics under S3 (i.e., site air temperature and the snowfall fraction) and S4 (with additional melting period; Figure 9d).

[40] For UADJ, the regression function of S4 is evidently different from its counterpart of S1 and S2, while the function of S3 only varies slightly from that of S1 and S2. This indicates that the impact of forcing uncertainty on UADJ is marginal, whereas the effect of structural uncertainty is significant. For DAYGM, the regression equations of S3 and S4 both vary from their respective counterparts of S1 and S2, illustrating that uncertainty in model forcing has a larger impact on DAYGM than on UADJ. The regression relationships for MBASE differ from each other under four scenarios. Particularly, the relationship associated with S4 is significantly different from that of the other three scenarios, in terms of both the number of independent site characteristics involved and the coefficients in the regression equation. This is also the case for parameter PXTEMP under S3 and S4. These observations highlight the complex influence of various uncertainty sources on parameter regionalization information. The influence is site dependent and parameter dependent and is thus difficult to generalize.

[41] To investigate the potential impact of various uncertainty scenarios on model parameters, scatter plots showing these four parameters derived from DREAM and the MLR regression equations at 24 study sites are presented (Figure 9). A clear feature observed for parameters UADJ, MBASE, and DAYGM is that parameter sets tend to cluster around the upper and lower bounds of the corresponding feasible parameter space under S1, S2, and S3. However, the parameters under S4 spread out across the entire feasible space. For instance, for DAYGM, the percentage of DAYGM values that are located within the range from 0.03 to 0.27 (the feasible parameter range for DAYGM is [0–0.3]; see Table 2) for S1, S2, S3, and S4 are 21%, 25%,

Table 4. Correlation Between DREAM-Derived Maximum Likelihood Parameters and Regression-Based Parameters^a

Parameters	S1		S2		S3		S4	
	<i>R</i>	<i>p</i> Value						
MFMAX	0.37	0.84	0.33	0.91	0.44	0.66	0.36	0.85
MFMIN	0.57	0.29	0.56	0.30	0.53	0.40	0.44	0.68
UADJ	<i>0.74</i>	0.02	<i>0.74</i>	0.02	0.52	0.41	0.36	0.85
MBASE	0.66	0.09	0.66	0.10	0.50	0.49	<i>0.72</i>	0.03
DAYGM	<i>0.72</i>	0.03	<i>0.73</i>	0.03	0.57	0.29	0.64	0.12
NMF	0.45	0.65	0.41	0.76	0.36	0.86	0.57	0.29
TIPM	0.45	0.65	0.41	0.76	0.47	0.58	0.39	0.79
PLWHC	0.34	0.89	0.33	0.90	0.26	0.97	0.32	0.92
SCF	-	-	-	-	0.64	0.13	0.57	0.28
PXTEMP	-	-	-	-	<i>0.74</i>	0.02	<i>0.72</i>	0.03

^aThe corresponding *p* values are also tabulated. The *p* values indicating significant correlation relationships are in bold; the associated correlation coefficient values are in italics.

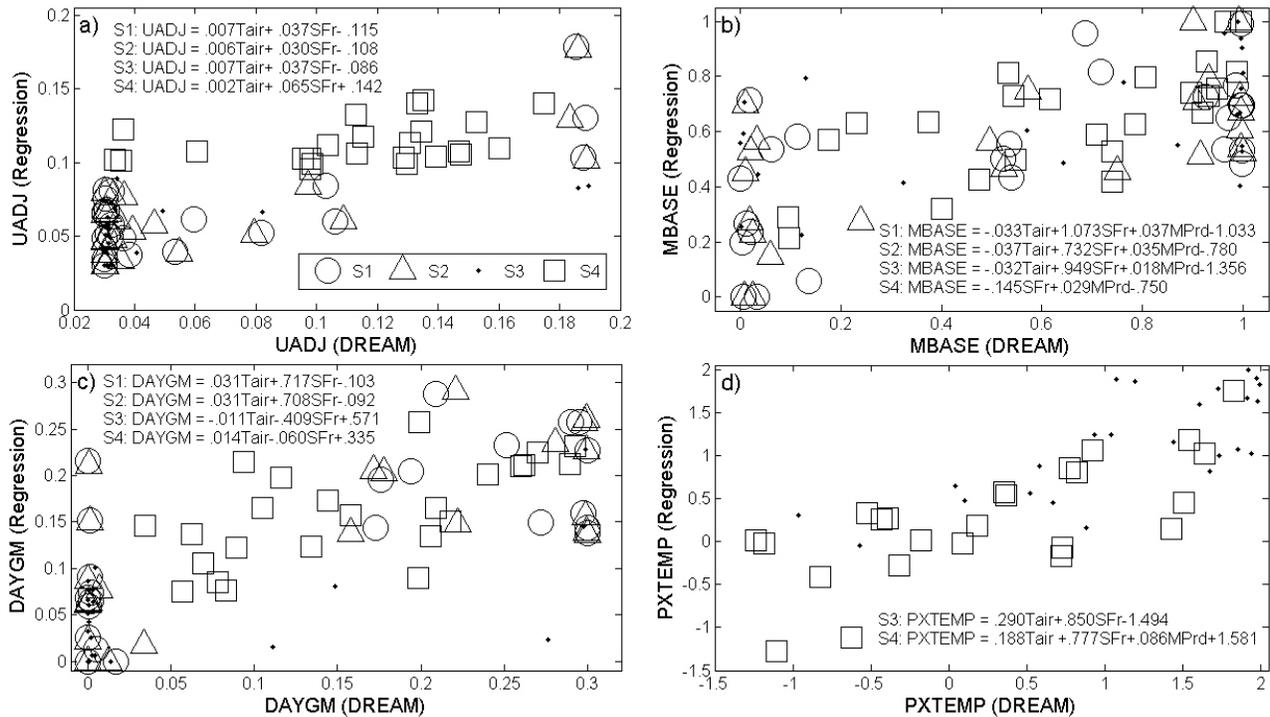


Figure 9. Scatterplot of DREAM-derived maximum likelihood parameter values (x axis) versus parameters derived from regression analysis (y axis) under various uncertainty scenarios for parameters (a) UADJ, (b) MBASE, (c) DAYGM, and (d) PXTEMP. Regression relationships are also shown. The site characteristics involved include mean annual air temperature (T_{air}), snowfall fraction (SFr), and the length of melting period in days (MPrd). PXTEMP is fixed in both S1 and S2 and thus is not included in Figure 9d.

13%, and 88%, respectively. PXTEMP in S3 also tends to cluster at the upper bound of the parameter range (Figure 9d), compared with that observed in S4.

5. Conclusions

[42] Regionalizing hydrologic model parameters holds great potential for improved hydrologic predictions in ungauged areas, particularly in light of a warming climate and the fact that the current hydrologic observational network across the United States is sparse. However, the presence of uncertainty in model forcing, parameters, structure, and measurements of model output (e.g., SWE) hinders the attempt to construct robust regionalization relationships for model parameters [Wagner and Wheat, 2006; Yadav et al., 2007; Bastola et al., 2008; Buytaert and Beven, 2009; Kling and Gupta, 2009]. The current study overviews the development of a comprehensive uncertainty analysis that can provide insight on the regionalization of hydrologic model parameters. The key findings of our study are summarized as follows:

[43] 1. The inclusion of uncertainty in SWE measurements (S2) results in little change in the 95% parameter uncertainty range, the marginal parameter distribution, the correlation among posterior parameters, the 95% SWE prediction uncertainty range, and the regionalization relationships of SNOW17 parameters. This information is almost identical to that derived when only parameter uncertainty (S1) is considered.

[44] 2. In contrast, the inclusion of uncertainty in snowfall forcing (S3) has a significant impact on the derived

parameter ranges and regionalization relationships. First, the posterior parameter uncertainty range changes in terms of both magnitude and location in the feasible parameter range. The wetter the sites, the more evident this change generally is. Second, the marginal distributions of most posterior parameters show differences in regard to either the type of distribution or the distribution characteristics (moments). Third, correlations among posterior parameters are dramatically different, with overall correlation becoming weaker with the addition of uncertainty in snowfall. Under S3, the SWE estimates at the wet sites studied are considerably improved compared to those estimated under cases S1 and S2. The regionalization relationships determined are also very different. For instance, parameter DAYGM is negatively correlated to site air temperature and the snowfall fraction, which is the opposite in the cases of S1 and S2.

[45] 3. Despite the evident consequences of uncertainty in snowfall forcing, model structural error (S4) tends to dominate. Explicit treatment of this uncertainty produces significantly wider parameter bounds, leading to wider ensembles of SWE predictions. In addition, when structural uncertainty is explicitly considered, there is no apparent marginal distribution of most posterior parameters at the three sites that were used for demonstration (uniform distributions are observed). The correlation among the posterior parameter samples is also weaker than the other three uncertainty scenarios at all the study sites. However, inclusion of model structural uncertainty does strengthen the overall correlation between parameter uncertainty and site characteristics. Furthermore, the regionalization relationships are dramatically

different from those obtained in S1, S2, and S3. MBASE and PXTMP are related to different site characteristics compared to their regionalization relationship in other scenarios. The maximum likelihood parameters and corresponding regression-based parameters spread out across the full feasible parameter space, while those parameters in S1, S2, and S3 tend to cluster around the upper and lower bounds of the feasible parameter space. As addressed in section 3.3.4, considering model structural uncertainty leads to more realistic estimates of model parameters. We advocate that the posterior parameters and corresponding SWE predictions as well as the regionalization information obtained under the most complex scenario (S4) are more realistic (and robust) than cases when structural uncertainty is not considered (i.e., S1, S2, and S3).

[46] While uncertainties in model forcing and model structure are shown to significantly alter regionalization relationships, it is difficult to rigorously characterize these changes. This may be due to several reasons. First, the specific site characteristics examined in this study may not be directly (or singularly) connected to the snow accumulation and ablation process [Merz and Blochl, 2004; Oudin et al., 2008]. Second, the relationships between SNOW17 parameters and site characteristics may be nonlinear, while the regression model applied in the current study assumes linearity. Third, a larger sample size may be needed [Young, 2006; Oudin et al., 2008]. In our ongoing work, which aims to develop and assess more robust regionalization schemes for SNOW17 parameters, we have started to investigate additional site characteristics (e.g., mean number of days when air temperature is close to 0°C) that may more closely correlate to the snow accumulation and ablation process. We are also investigating the use of nonlinear regression techniques (i.e., ANNs) to maximize the chance of finding useful relationships between model parameters and site characteristics. The source code of DREAM can be obtained from the last author upon request.

[47] **Acknowledgments.** We would like to thank Pedro Restrepo, Dong-June Seo, and Yuqiong Liu for their suggestions and advice on the development of this work. This research was partially supported by a grant from the National Oceanic Atmospheric Administration (NOAA) National Weather Service (NA07NWS4620013) and by a UCLA Graduate Division Fellowship. The work of the last author was supported by a J. Robert Oppenheimer Fellowship from the Los Alamos National Laboratory Postdoctoral Program. We would also like to thank Newsha Ajami and two anonymous reviewers for their constructive comments on an earlier version of this paper.

References

- Ajami, N. K., Q. Duan, and S. Sorooshian (2007), An integrated hydrologic Bayesian multimodel combination framework: Confronting input, parameter, and model structural uncertainty in hydrologic prediction, *Water Resour. Res.*, *43*, W01403, doi:10.1029/2005WR004745.
- Anderson, E. A. (1973), National Weather Service river forecast system—Snow accumulation and ablation model, *Tech. Memo NWS HYDRO-17*, NOAA, Silver Spring, Md.
- Anderson, E. A. (2002), Calibration of conceptual models for use in river forecasting, *NOAA Tech. Rep. NWS 45*, Hydro. Lab., Silver Spring, Md.
- Andreadis, K. M., and D. P. Lettenmaier (2006), Assimilating remotely sensed snow observations into a macroscale hydrology model, *Adv. Water Resour.*, *29*, 872–886.
- Bales, R. C., N. P. Molotch, T. H. Painter, M. D. Dettinger, R. Rice, and J. Dozier (2006), Mountain hydrology of the western United States, *Water Resour. Res.*, *42*, W08432, doi:10.1029/2005WR004387.
- Bastola, S., H. Ishidaira, and K. Takeuchi (2008), Regionalisation of hydrological model parameters under parameter uncertainty: A case study involving TOPMODEL and basins across the globe, *J. Hydrol.*, *357*, 188–206.
- Beck, M. B. (1987), Water quality modeling: A review of the analysis of uncertainty, *Water Resour. Res.*, *23*, 1393–1442, doi:10.1029/WR023i008p01393.
- Beven, K. (2006), A manifesto for the equifinality thesis, *J. Hydrol.*, *320*, 18–36.
- Beven, K., and A. Binley (1992), Future of distributed models: Model calibration and uncertainty prediction, *Hydrol. Processes*, *6*, 279–298.
- Buytaert, W., and K. Beven (2009), Regionalization as a learning process, *Water Resour. Res.*, *45*, W11419, doi:10.1029/2008WR007359.
- Campbell, E. P., and B. C. Bates (2001), Regionalization of rainfall-runoff model parameters using Markov chain Monte Carlo samples, *Water Resour. Res.*, *37*, 731–739, doi:10.1029/2000WR900349.
- Cayan, D. R. (1996), Interannual climate variability and snowpack in the western United States, *J. Clim.*, *9*, 928–948.
- Cayan, D. R., M. D. Dettinger, S. A. Kammerdiener, J. M. Caprio, and D. H. Peterson (2001), Changes in the onset of spring in the western United States, *Bull. Am. Meteorol. Soc.*, *82*, 399–415.
- Clark, M. P., and D. Kavetski (2010), Ancient numerical daemons of conceptual hydrological modeling: 1. Fidelity and efficiency of time stepping schemes, *Water Resour. Res.*, *46*, W10510, doi:10.1029/2009WR008894.
- Clark, M. P., and A. G. Slater (2006), Probabilistic quantitative precipitation estimation in complex terrain, *J. Hydrometeorol.*, *7*, 3–22.
- Clark, M. P., A. G. Slater, D. E. Rupp, R. A. Woods, J. A. Vrugt, H. V. Gupta, T. Wagener, and L. E. Hay (2008), Framework for Understanding Structural Errors (FUSE): A modular framework to diagnose differences between hydrological models, *Water Resour. Res.*, *44*, W00B02, doi:10.1029/2007WR006735.
- Crook, A. G. (1977), SNOTEL: Monitoring climatic factors to predict water supplies, *J. Soil Water Conserv.*, *32*, 294–295.
- Doesken, N. J., and G. L. Schaefer (1987), The contribution of SNOTEL precipitation measurements to climate analysis, monitoring and research, paper presented at Western Snow Conference, Vancouver, B. C., Canada, 14–16 April.
- Duan, Q., S. Sorooshian, and V. Gupta (1992), Effective and efficient global optimization for conceptual rainfall-runoff models, *Water Resour. Res.*, *28*, 1015–1031, doi:10.1029/91WR02985.
- Duan, Q., N. K. Ajami, X. Gao, and S. Sorooshian (2007), Multi-model ensemble hydrologic prediction using Bayesian model averaging, *Adv. Water Resour.*, *30*, 1371–1386.
- Durand, M., and S. A. Margulis (2006), Feasibility test of multifrequency radiometric data assimilation to estimate snow water equivalent, *J. Hydrometeorol.*, *7*, 443–457.
- Durand, M., and S. A. Margulis (2008), Effects of uncertainty magnitude and accuracy on assimilation of multiscale measurements for snowpack characterization, *J. Geophys. Res.*, *113*, D02105, doi:10.1029/2007JD008662.
- Fassnacht, S. R., K. A. Dressler, and R. C. Bales (2003), Snow water equivalent interpolation for the Colorado River Basin from snow telemetry (SNOTEL) data, *Water Resour. Res.*, *39*(8), 1208, doi:10.1029/2002WR001512.
- Franz, K. J. (2006), Characterization of the comparative skill of conceptual and physically-based snow models for streamflow prediction, Ph.D. thesis, Univ. of Calif., Irvine.
- Franz, K. J., T. S. Hogue, and S. Sorooshian (2008), Operational snow modeling: Addressing the challenges of an energy balance model for National Weather Service forecasts, *J. Hydrol.*, *360*, 48–66.
- Franz, K. J., P. Butcher, and N. K. Ajami (2010), Addressing snow model uncertainty for hydrologic prediction, *Adv. Water Resour.*, *33*, 820–832.
- Gelman, A., and D. B. Rubin (1992), Inference from iterative simulation using multiple sequences, *Stat. Sci.*, *7*, 457–472.
- Gotzinger, J., and A. Bardossy (2008), Generic error model for calibration and uncertainty estimation of hydrological models, *Water Resour. Res.*, *44*, W00B07, doi:10.1029/2007WR006691.
- Gupta, H. V., K. J. Beven, and T. Wagener (2006), Model calibration and uncertainty estimation, in *Encyclopedia of Hydrologic Sciences*, edited by M. G. Anderson and J. J. McConnell, pp. 2015–2031, John Wiley, New York.
- He, M. (2010), Data assimilation in watershed models for improved hydrologic forecasting, 173 pp, University of California, Los Angeles.
- He, M., T. S. Hogue, K. J. Franz, S. A. Margulis, and J. A. Vrugt (2011), Characterizing parameter sensitivity and uncertainty for a snow model across hydroclimatic regimes, *Adv. Water Resour.*, *34*, 114–127.

- Heuvelmans, G., B. Muys, and J. Feyen (2006), Regionalisation of the parameters of a hydrological model: Comparison of linear regression models with artificial neural nets, *J. Hydrol.*, *319*, 245–265.
- Hogue, T. S., S. Sorooshian, H. Gupta, A. Holz, and D. Braatz (2000), A multistep automatic calibration scheme for river forecasting models, *J. Hydrometeorol.*, *1*, 524–542.
- Hogue, T. S., L. A. Bastidas, H. V. Gupta, and S. Sorooshian (2006a), Evaluating model performance and parameter behavior for varying levels of land surface model complexity, *Water Resour. Res.*, *42*, W08430, doi:10.1029/2005WR004440.
- Hogue, T. S., H. Gupta, and S. Sorooshian (2006b), A ‘user-friendly’ approach to parameter estimation in hydrologic models, *J. Hydrol.*, *320*, 202–217.
- Huisman, J. A., J. Rings, J. A. Vrugt, J. Sorg, and H. Vereecken (2010), Hydraulic properties of a model dike from coupled Bayesian and multi-criteria hydrogeophysical inversion, *J. Hydrol.*, *380*, 62–73.
- Jin, J., X. Gao, S. Sorooshian, Z. L. Yang, R. Bales, R. E. Dickinson, S. F. Sun, and G. X. Wu (1999), One-dimensional snow water and energy balance model for vegetated surfaces, *Hydrol. Processes*, *13*, 2467–2482.
- Kapnick, S., and A. Hall (2010), Observed climate-snowpack relationships in California and their implications for the future, *J. Clim.*, *23*, 3446–3456.
- Kavetski, D., and M. P. Clark (2010), Ancient numerical demons of conceptual hydrological modeling: 2. Impact of time stepping schemes on model analysis and prediction, *Water Resour. Res.*, *46*, W10511, doi:10.1029/2009WR008896.
- Kavetski, D., S. W. Franks, and G. Kuczera (2002), Confronting input uncertainty in environmental modelling, in *Calibration of Watershed Models*, *Water Sci. Appl. Ser.*, vol. 6, edited by Q. Duan et al., pp. 49–68, AGU, Washington, D. C.
- Kavetski, D., G. Kuczera, and S. W. Franks (2006), Bayesian analysis of input uncertainty in hydrological modeling: 1. Theory, *Water Resour. Res.*, *42*, W03407, doi:10.1029/2005WR004368.
- Kitanidis, P. K., and R. L. Bras (1980), Real-time forecasting with a conceptual hydrologic model: 1. Analysis of uncertainty, *Water Resour. Res.*, *16*, 1025–1033, doi:10.1029/WR016i006p01025.
- Kling, H., and H. Gupta (2009), On the development of regionalization relationships for lumped watershed models: The impact of ignoring sub-basin scale variability, *J. Hydrol.*, *373*, 337–351.
- Kuczera, G., and E. Parent (1998), Monte Carlo assessment of parameter uncertainty in conceptual catchment models: The Metropolis algorithm, *J. Hydrol.*, *211*, 69–85.
- Kuczera, G., D. Kavetski, S. Franks, and M. Thyer (2006), Towards a Bayesian total error analysis of conceptual rainfall-runoff models: Characterising model error using storm-dependent parameters, *J. Hydrol.*, *331*, 161–177.
- Leclerc, M., and T. B. M. J. Ouarda (2007), Non-stationary regional flood frequency analysis at ungauged sites, *J. Hydrol.*, *343*, 254–265.
- Liu, Y., and H. V. Gupta (2007), Uncertainty in hydrologic modeling: Toward an integrated data assimilation framework, *Water Resour. Res.*, *43*, W07401, doi:10.1029/2006WR005756.
- Margulis, S. A., D. McLaughlin, D. Entekhabi, and S. Dunne (2002), Land data assimilation and estimation of soil moisture using measurements from the Southern Great Plains 1997 Field Experiment, *Water Resour. Res.*, *38*(12), 1299, doi:10.1029/2001WR001114.
- Merz, R., and G. Blöschl (2004), Regionalisation of catchment model parameters, *J. Hydrol.*, *287*, 95–123.
- Mock, C. J. (1996), Climatic controls and spatial variations of precipitation in the western United States, *J. Clim.*, *9*, 1111–1125.
- Molotch, N. P., and R. C. Bales (2005), Scaling snow observations from the point to the grid element: Implications for observation network design, *Water Resour. Res.*, *41*, W11421, doi:10.1029/2005WR004229.
- Montanari, A. (2007), What do we mean by ‘uncertainty’? The need for a consistent wording about uncertainty assessment in hydrology, *Hydrol. Processes*, *21*, 841–845.
- Mote, P. W. (2003), Trends in snow water equivalent in the Pacific Northwest and their climatic causes, *Geophys. Res. Lett.*, *30*(12), 1601, doi:10.1029/2003GL017258.
- Mote, P. W., A. F. Hamlet, M. P. Clark, and D. P. Lettenmaier (2005), Declining mountain snowpack in western North America, *Bull. Am. Meteorol. Soc.*, *86*, 39–49.
- Muleta, M. K., and J. W. Nicklow (2005), Sensitivity and uncertainty analysis coupled with automatic calibration for a distributed watershed model, *J. Hydrol.*, *306*, 127–145.
- National Weather Service (NWS) (2004), National Weather Service River Forecast System—User’s manual, Silver Spring, Md.
- Oudin, L., V. Andréassian, C. Perrin, C. Michel, and N. Le Moine (2008), Spatial proximity, physical similarity, regression and ungauged catchments: A comparison of regionalization approaches based on 913 French catchments, *Water Resour. Res.*, *44*, W03413, doi:10.1029/2007WR006240.
- Parajka, J., R. Merz, and G. Blöschl (2005), A comparison of regionalisation methods for catchment model parameters, *Hydrol. Earth Syst. Sci.*, *9*, 157–171.
- Renard, B., D. Kavetski, G. Kuczera, M. Thyer, and S. W. Franks (2010), Understanding predictive uncertainty in hydrologic modeling: The challenge of identifying input and structural errors, *Water Resour. Res.*, *46*, W05521, doi:10.1029/2009WR008328.
- Restrepo, P. J., and R. L. Bras (1985), A view of maximum-likelihood estimation with large conceptual hydrologic models, *Appl. Math. Comput.*, *17*, 375–403.
- Salamon, P., and L. Feyen (2009), Assessing parameter, precipitation, and predictive uncertainty in a distributed hydrological model using sequential data assimilation with the particle filter, *J. Hydrol.*, *376*, 428–442.
- Schoups, G., J. A. Vrugt, F. Fenicia, and N. C. van de Giesen (2010), Corruption of accuracy and efficiency of Markov chain Monte Carlo simulation by inaccurate numerical implementation of conceptual hydrologic models, *Water Resour. Res.*, *46*, W10530, doi:10.1029/2009WR008648.
- Seidou, O., and T. B. M. J. Ouarda (2007), Recursion-based multiple changepoint detection in multiple linear regression and application to river streamflows, *Water Resour. Res.*, *43*, W07404, doi:10.1029/2006WR005021.
- Serreze, M. C., M. P. Clark, R. L. Armstrong, D. A. McGinnis, and R. S. Pulwarty (1999), Characteristics of the western United States snowpack from snowpack telemetry (SNOTEL) data, *Water Resour. Res.*, *35*, 2145–2160, doi:10.1029/2000WR900307.
- Serreze, M. C., M. P. Clark, and A. Frei (2001), Characteristics of large snowfall events in the montane western United States as examined using snowpack telemetry (SNOTEL) data, *Water Resour. Res.*, *37*, 675–688, doi:10.1029/2000WR900307.
- Service, R. F. (2004), As the west goes dry, *Science*, *303*, 1124–1127.
- Slater, A. G., and M. P. Clark (2006), Snow data assimilation via an ensemble Kalman filter, *J. Hydrometeorol.*, *7*, 478–493.
- Sorooshian, S., and J. A. Dracup (1980), Stochastic parameter estimation procedures for hydrologic rainfall-runoff models: Correlated and heteroscedastic error cases, *Water Resour. Res.*, *16*, 430–442, doi:10.1029/WR016i002p00430.
- Stewart, I. T., D. R. Cayan, and M. D. Dettinger (2004), Changes in snowmelt runoff timing in western North America under a ‘business as usual’ climate change scenario, *Clim. Change*, *62*, 217–232.
- Strasser, U., P. Etchevers, and Y. Lejeune (2002), Inter-comparison of two snow models with different complexity using data from an alpine site, *Nord. Hydrol.*, *33*, 15–26.
- Tarboton, D. G., and C. H. Luce (1996), Utah Energy Balance Snow Accumulation and Melt Model (UEB): Computer model technical description and users guide, Utah Water Res. Lab. Logan.
- Thiemann, M., M. Trosset, H. Gupta, and S. Sorooshian (2001), Bayesian recursive parameter estimation for hydrologic models, *Water Resour. Res.*, *37*, 2521–2535, doi:10.1029/2000WR900405.
- Thyer, M., B. Renard, D. Kavetski, G. Kuczera, S. W. Franks, and S. Srikanthan (2009), Critical evaluation of parameter consistency and predictive uncertainty in hydrological modeling: A case study using Bayesian total error analysis, *Water Resour. Res.*, *45*, W00B14, doi:10.1029/2008WR006825.
- Tiwari, M. K., and C. Chatterjee (2010), Development of an accurate and reliable hourly flood forecasting model using wavelet-bootstrap-ANN (WBANN) hybrid approach, *J. Hydrol.*, *394*, 458–470.
- Tribbeck, M. J., R. J. Gurney, E. M. Morris, and D. W. C. Pearson (2004), A new snow-SVAT to simulate the accumulation and ablation of seasonal snow cover beneath a forest canopy, *J. Glaciol.*, *50*, 171–182.
- Vrugt, J. A., and B. A. Robinson (2007), Treatment of uncertainty using ensemble methods: Comparison of sequential data assimilation and Bayesian model averaging, *Water Resour. Res.*, *43*, W01411, doi:10.1029/2005WR004838.
- Vrugt, J. A., W. Bouten, H. V. Gupta, and S. Sorooshian (2002), Toward improved identifiability of hydrologic model parameters: The information content of experimental data, *Water Resour. Res.*, *38*(12), 1312, doi:10.1029/2001WR001118.
- Vrugt, J. A., H. V. Gupta, W. Bouten, and S. Sorooshian (2003), A shuffled complex evolution Metropolis algorithm for optimization and uncertainty assessment of hydrological model parameters, *Water Resour. Res.*, *39*(8), 1201, doi:10.1029/2002WR001642.

- Vrugt, J. A., C. G. H. Diks, H. V. Gupta, W. Bouten, and J. M. Verstraten (2005), Improved treatment of uncertainty in hydrologic modeling: Combining the strengths of global optimization and data assimilation, *Water Resour. Res.*, *41*, W01017, doi:10.1029/2004WR003059.
- Vrugt, J. A., C. ter Braak, H. Gupta, and D. Robertson (2008a), Equifinality of formal (DREAM) and informal (GLUE) Bayesian approaches in hydrologic modeling?, *Stochastic Environ. Res. Risk Assess.*, *23*, 1011–1026, doi:10.1007/s00477-008-0274-y.
- Vrugt, J. A., C. J. F. ter Braak, M. P. Clark, J. M. Hyman, and B. A. Robinson (2008b), Treatment of input uncertainty in hydrologic modeling: Doing hydrology backward with Markov chain Monte Carlo simulation, *Water Resour. Res.*, *44*, W00B09, doi:10.1029/2007WR006720.
- Vrugt, J. A., C. ter Braak, C. G. H. Diks, D. A. Roberts, J. M. Hyman, and D. Higdon (2009), Accelerating Markov chain Monte Carlo simulation by differential evolution with self-adaptive randomized subspace sampling, *Int. J. Nonlinear Sci. Numer. Simul.*, *10*, 273–290.
- Wagener, T., and H. S. Wheater (2006), Parameter estimation and regionalization for continuous rainfall-runoff models including uncertainty, *J. Hydrol.*, *320*, 132–154.
- Wagener, T., N. McIntyre, M. J. Lees, H. S. Wheater, and H. V. Gupta (2003), Towards reduced uncertainty in conceptual rainfall-runoff modelling: Dynamic identifiability analysis, *Hydrol. Processes*, *17*, 455–476.
- Yadav, M., T. Wagener, and H. Gupta (2007), Regionalization of constraints on expected watershed response behavior for improved predictions in ungauged basins, *Adv. Water Resour.*, *30*, 1756–1774.
- Young, A. R. (2006), Stream flow simulation within UK ungauged catchments using a daily rainfall-runoff model, *J. Hydrol.*, *320*, 155–172.
-
- K. J. Franz, Department of Geological and Atmospheric Sciences, Iowa State University, Ames, IA 50011, USA.
- M. He, Office of Hydrologic Development, National Weather Service, NOAA, 1325 East West Hwy., Silver Spring, MD 20910, USA.
- T. S. Hogue and S. A. Margulis, Department of Civil and Environmental Engineering, University of California, 5731F Boelter Hall, Los Angeles, CA 90095-1593, USA. (thogue@seas.ucla.edu)
- J. A. Vrugt, Department of Civil and Environmental Engineering, University of California, Irvine, CA 92697, USA.